Testing of Vote Recorders

Final Report
August 24, 2001

Michael Hites, Ph.D
Bill Ornt

Illinois Institute of Technology

Office of Technology Services
10 West 31st Street, Room 007
Chicago, IL 60616
and
Department of Mechanical, Materials and Aerospace Engineering
10 West 32nd Street, Room 243
Chicago, IL 60616and

hites@iit.edu

Summary

Illinois Institute of Technology tested 219 vote recorders supplied by the Chicago Board of Election Commissioners and used in the November 2000 election. The vote recorders were chosen based on a combination of two selection processes. One-hundred nineteen of the vote recorders were chosen by the Chicago Board of Election Commissioners as possible "suspect" vote recorders, while the remaining 100 were chosen randomly, two from each ward. These vote recorders were inspected, categorized and tested to determine if there were any systematic, repeatable errors associated with the voting hardware itself. The goal was to test each vote recorder in a similar environment to an election, but perform specific tests to determine the accuracy of the machine to record a correct vote. This is a separate issue from usability of the device, and usability is not a subject of this study.

Each vote recorder contains one of two templates, denoted here a Template 1 and Template 2. The template is used to guide the voting stylus into the ballot card. During the template manufacturing process, the two templates are made in simultaneously from a single mold comprised of two cavities, so that approximately half of the vote recorders use Template 1 and the other half Template 2. In our sample of 219 vote recorders, approximately 60% were Template 1 and 40% Template 2. In the subset of 119 suspect vote recorders, about 70% were Template 1, and in the subset of 100 random vote recorders, about 55% were Template 1.

The measurements presented here indicate there is a high degree of correlation between high error rates in vote recorders and the use of Template 1. This conclusion is based on the following observations:

- The problem with the misalignment of Template 1 was the major contributor to the voting errors that occurred during our testing.
- None of the other vote recorder defects, other than the manufacturing defects in the templates, contributed to high error rates.
- The probability of a vote recorder with a high error rate having Template 1 is about 20% higher than expected.
- Tests with vote recorders using Template 2 consistently had error rates less than 1%.
- Tests with vote recorders using Template 1 showed an average error rate of about 4% (in Test 7).
- Measurements of the template manufacturing accuracy showed that both templates were out of tolerance, however the holes in Template 1 were lower than specified by twice the amount of Template 2.
- Each of the voters in the study showed an increase in error rates when comparing tests with Template 2 to tests with Template 1.
- A gross summary of all tests with Template 1 vote recorders showed an average of 2.5% error, while Template 2 vote recorders showed an average of 0.3% error.
- Statistical analysis shows that the odds of getting an unsuccessful vote is 9.1 times more likely with Template 1 than Template 2 at a confidence level of greater than 99%.

The recommendation is to remanufacture the templates used in the vote recorders. Similar tests should be repeated with a representative set of new vote recorders to rule out the possibility of equipment defects as a source of voter error in future elections.

Tests Conducted

The vote recorder of interest is the Votomatic with 456-position punch card. The recorder is used with the Poll Star portable voting booth. For all tests performed here, the card was inserted into the data punch, and the manufacturer's stylus was used to punch the card. The punched cards were inspected visually, and the results were tabulated manually after testing. The condition of each punch was noted and recorded according to Table 1 for each punch. For the purpose of reporting the error rate, any punch that resulted in a "zero", "two" or "three" from Table 1 was considered an error.

Table 1 – Parameters for recording results of punched cards

Recorded Value	**Corresponding Condition**
0	No evidence of voter intent
1	Correctly punched chad
2	Hanging chad
3	Dimpled chad

Two-hundred nineteen vote recorders were used in this study. The vote recorders were delivered to Illinois Institute of Technology by the Chicago Board of Election Commissioners in 11 boxes. The vote recorders were removed from the boxes, and each vote recorder was given a unique number ranging from IIT001 to IIT219. Each vote recorder was visually inspected and the physical defects were noted and recorder. The attachment "Vote Recorder Inventory 070201" details the vote recorder number, box number, and physical attributes of each vote recorder used in these tests. It was noted that 60% of the vote recorders used template 1 and 40% used template 2.

In all of the tests, the vote recorders were placed in standard Poll Star voting booths (used in the November 2000 election) and were setup in a climate-controlled room in the Stuart Building at the Illinois Institute of Technology Main Campus.

The voters in all of the tests were staff and students at Illinois Institute of Technology. Each voter was assigned a voter number according to Table 2.

Table 2 – Voters used in testing

Voter Number	**Role at IIT**
01	Faculty
02	Staff
03	Graduate Student
04	Undergraduate Student
05	Undergraduate Student
06	Undergraduate Student
07	Undergraduate Student
08	Undergraduate Student

Following are the details of the tests conducted.

Test 1
Diagnostic test with originally configured vote recorders
This test was used to assess the initial status of the vote recorder as they were received. The test was performed before disassembling the vote recorder for visual inspection of defects.

- Each of the 219 vote recorders were removed from the shipping containers and tested as received from the Chicago Board of Election Commissioners. Each of the recorders was tested as used in the November 2000 election.
- All available ballot holes (191 ±2 total) were punched in the ballot book.
- One ballot per machine was tested using five voters.

Test 2
Standard diagnostic test: all 456 holes punched
This test was performed to duplicate the validation punch test referenced by Exhibit E, "Acceptance Testing Specifications" as noted in the December 14, 1999 Agreement of Purchase of Vote Recorders Between Election Works Corporation and Board of Election Commissioners of the City of Chicago.

- Each of the 219 vote recorders was tested with the ballot book removed.
- All 456 holes were punched for each card.
- One ballot per machine was tested using five voters.
- Since approximately 100,000 individual punches were performed, this was the only test performed with the students in both the standing and sitting position to eliminate voter fatigue.

Test 3
Diagnostic test focusing on presidential punches
This test focused on the holes with the most relevance and the highest error rates in the November 2000 election, namely the presidential race. The number of punched holes was reduced from 191 in Test 1 and 456 in Test 2, to 18 in Test 3 so that a greater number of cards could be punched in later tests.

- Forty-one vote recorders were tested with the ballot book in place.
- Eighteen standard punches were developed based on the position of the presidential election holes in the template. The locations were selected to include the presidential election column along with holes in other, lesser punched, columns were holes could be selected in the same row as the presidential holes.
- Exact holes punched were 2, 4, 6, 8, 10, 12, 230, 232, 234, 236, 238, 240, 306, 308, 310, 312, 314, 316.
- One ballot per machine was tested using five voters.

Test 4
Retest of machines in Test 3
This test was used to gather a greater number of punched ballots to obtain enough data for statistical relevance.

- Thirty-eight machines from Test 3 were retested to increase the number of cards punched for the machines with high error rates.
- The same eighteen standard holes from Test 3 were punched.
- Exact holes punched were 2, 4, 6, 8, 10, 12, 230, 232, 234, 236, 238, 240, 306, 308,

310, 312, 314, 316.
- Twelve ballots per machine were tested using four voters.

Test 5
Detailed test of "normal" machines with few errors
This test was use to document the behavior of machines that showed error rates consistently less than 1%. 100 ballots were used for each machine to decrease statistical uncertainty.
- Twenty-five machines with no apparent defects or errors were used
- Test was performed with 60% template 1, 40% template 2 which is consistent with the overall ratio of the templates received.
- The same eighteen standard holes from Test 3 were punched.
- Exact holes punched were 2, 4, 6, 8, 10, 12, 230, 232, 234, 236, 238, 240, 306, 308, 310, 312, 314, 316.
- Twenty-five ballots per machine were tested using four voters (100 ballots per machine)

Test 6
Test to determine probably of voting success with template misalignment
This test was performed to estimate the misalignment of the template and the ballot book that would result in an unacceptable error rate (5% or greater). Since the template and the ballot move in tandem when the card is inserted, it was believed that it was the misalignment of the template holes with the ballot book holes could be a major contributor in the error rate of the vote recorder. The vote recorder that was used in this test contained Template 2 and was machined (before it was delivered to IIT) so that it was within the manufacturer's specifications.
- A single vote recorder was used in this test (IIT 028).
- The same eighteen standard holes from Test 3 were punched.
- Exact holes punched were 2, 4, 6, 8, 10, 12, 230, 232, 234, 236, 238, 240, 306, 308, 310, 312, 314, 316.
- One-hundred ballots were tested using two student voters.
- The one-hundred ballots were inserted into the vote recorders at various to various lengths by elongating the ballot pin holes so from the original length of 0.315" to 0.400" in small increments.
- Voting error was plotted versus linear translation from the normal stopping position of the template when the card is inserted.

Test 7 (repeat of test 5 with different machines)
Detailed test of "suspect" machines with high errors
This test was use to document the behavior of machines that showed errors consistently in tests 1, 2, 3, and 4. It was concluded that machines that showed more than two instanances of high error rates in two of the four tests should be considered for this group to represent the worst of the vote recorders tested. 100 ballots were used for each machine to decrease statistical uncertainty.
- Twenty-five machines with high error rates were used.
- Test was performed with 60% template 1, 40% template 2 which is consistent with the overall ratio of the templates received.
- The same eighteen standard holes from Test 3 were punched.
- Exact holes punched were 2, 4, 6, 8, 10, 12, 230, 232, 234, 236, 238, 240, 306, 308,

310, 312, 314, 316.
- Twenty-five ballots per machine were tested using four voters (100 ballots per machine).

Test 8
Measurement of template hole location for discrepancy in manufactured dimensions for "normal" templates
- Seven templates from the vote recorders used in Test 5 were measured using a Mitutoyo Quick Scope non-contact optical measuring system.
- Images were obtained for eight holes on each of seven templates for a total of 56 measurements and compared to the original template drawings that were submitted to the manufacturer for fabrication.

Test 9
Measurement of template hole location for discrepancy in manufactured dimensions for "suspect" templates
- Six templates from the vote recorders used in Test 7 were measured using a Mitutoyo Quick Scope non-contact optical measuring system.
- Images were obtained for eight holes on each of six templates for a total of 48 measurements and compared to the original template drawings that were submitted to the manufacturer for fabrication.

Results

Visual inspection of the vote recorders was not as important as testing the performance of the vote recorders by simulating an election, however there are a few important statements to be made about the condition of the vote recorders as received. First, three of 219 had no templates. If these were used in the voting process, they inevitably would have caused bad ballots. Secondly, five of the ballot books had all of the holes punched, not just the ones where there were names of candidates on the ballots. If these were used in the election, they might lead to voter confusion, but probably not additional errors. Additionally, it was observed that the type of spring used in the template (plastic vs. metal) made no apparent difference in the performance of the recorder, however the metal springs appeared to be more likely loosen and "wiggle" than the plastic springs. This did not appear to affect performance.

In Test 1, 219 vote recorders were examined by punching every hole in the November 2000 ballot with one ballot card per vote recorder. By counting all hanging chads, dimpled chads and no-punches as errors, the tested showed that 40 out of 219 vote recorders (18%) had error rates of 2% or greater, 23 out of 219 (10%) had error rates of 5% or higher and, 14 out of 219 (6%) had error rates of 10% or higher.

Each vote recorder contains one of two templates, denoted here a Template 1 and Template 2. During the template manufacturing process, the two templates are made simultaneously from the single mold comprised of two cavities, so that approximately half of the vote recorders use Template 1 and the other half Template 2. In our sample of 219 vote recorders, 60% were Template 1 and 40% Template 2.

Figure 1 represents the probability that the vote recorder contains a Template 1 as a function of the error rate measured from the 219 vote recorders in Test 1. The horizontal axis is the percentile of the highest error rate, which in this figure means that the lowest percentile (10%) represents the vote recorders that had the lowest errors, while the highest percentile (90%) represents the vote recorders that had the highest error rates. The probability was calculated by sorting the error rates for each vote recorder from highest to lowest and then taking a running average of the probability that 20 consecutive vote recorders contained a template 1. A line is drawn at 60% to represent the theoretical probability of containing a Template 1 if the two templates were manufactured identically and to specification. Figure 1 clearly shows that the probability of the vote recorder containing Template 1 is far above the 60% theoretical line for the top 20% of vote recorders with the highest errors. Based on Figure 1, the chance of a vote recorder with a high error rate having Template 1 is about 80%, when it should be about 60%.

In Test 2, where all 456 holes were punched in one ballot card for each machine, the ballot book was removed so that the template could easily be seen. Because of this, the error rates were expectedly lower. In Test 2, only 14 vote recorders (6%) showed error rates of greater than 1% and only 9 vote recorders (4%) greater than 2%.

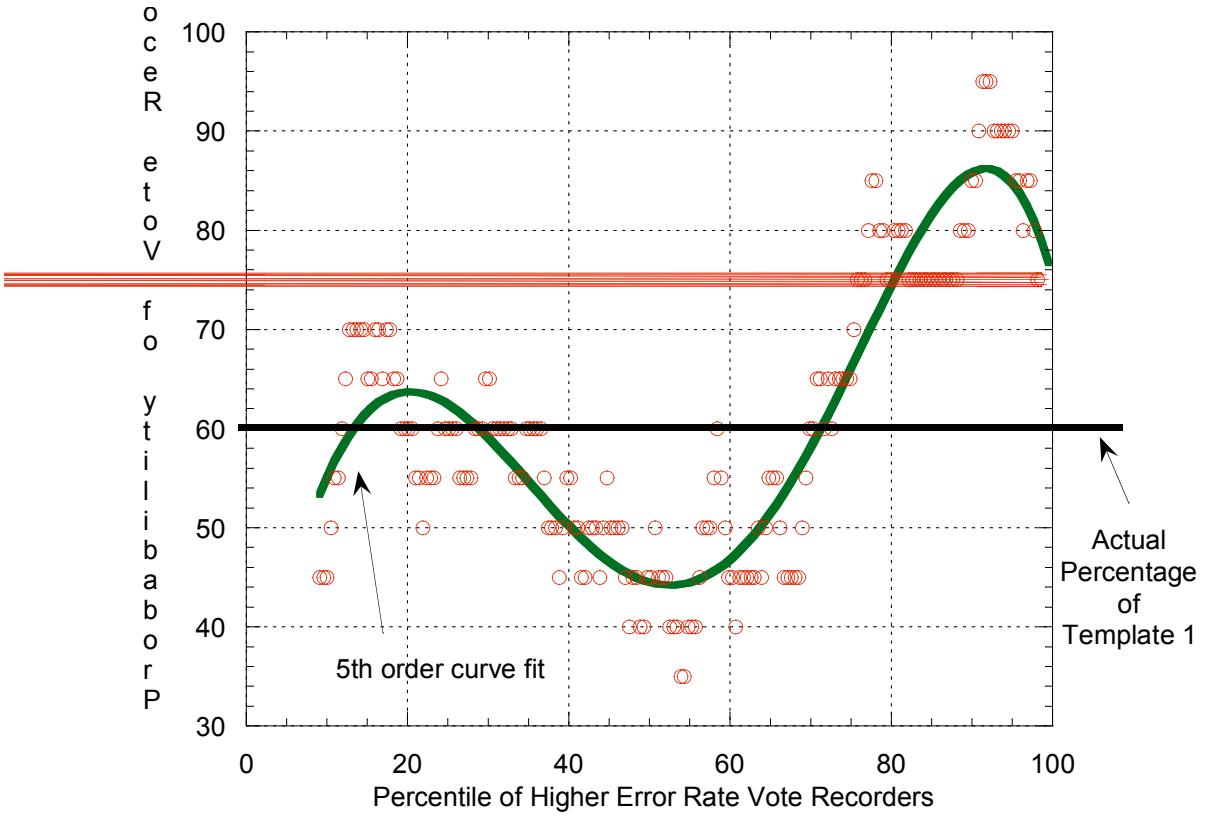

Figure 1 – Probability of a tested vote recorder containing Template 1 as a function of the error rate.

In Test 3, the number of holes punched was reduced from 191 in Test 1 to eighteen so that a greater number of cards could be punched to increase statistical accuracy. The eighteen holes were chosen to include the presidential election, where the highest error was observed in the November 2000 election, and two other columns were chosen within the ballot book with "lesser-punched" elections or retention votes. The number of vote recorders was reduced from 219 to 41 and included vote recorders with both high and low errors. In Test 4, the same test was repeated with 38 of the 41 vote recorders from Test 3. Between Tests 3 and 4, a total of 13 cards were punched for each of the machines using four different voters. Test 3 and 4 were used primarily to select the vote recorders that produced the highest errors so that this could be used as the control group for Test 7 and compared to the low-error vote recorders in Test 5.

In addition, Test 3 and 4 revealed that of the 18 holes punched during the tests, the number-four hole (presidential candidate Al Gore) consistently had the highest error rate. As of yet, this has not been explained, but it appears to be correlated with template number. Subsequent measurement of the hole diameter showed that the number four hole was about 0.001" larger in diameter than the number three hole that was unused for the November 2000 election and about 0.0005" larger in diameter than the number two (presidential candidate George Bush). Additional investigation could be performed to determine if wear-and-tear on the vote recorder is a factor in the error rate, however in Test 5, which examined the vote recorders with low error rates, there were no holes with substantially higher error rates than the others. The comparison of Test 5 to Test 7 is the most revealing comparison of the study. Test 5 represented the "normal" vote recorders in that they had characteristically low error rates.

Test 7 represented the "suspect" vote recorders in that they had characteristically higher error rates. Test 7 vote recorders were 009, 013, 015, 016, 029, 058, 059, 062, 063, 071, 072, 084, 128, 130, 135, 141, 147, 182 and 187.

In Test 5, the ratio of Template 1 to Template 2 was maintained at 60% to 40%, respectively. The colors denote the relative error. It can be seen from Table 3 that the vote recorders that contained Template 1 generally had higher or much higher error rates than Template 2. All of the vote recorders with Template 2 had error rates less than 1%, while all of the vote recorders with error rates greater than 1% used Template 1. The overall error rate for all machines in this test was about 1.5%, which seems rather high considering that this test represented a subset of "normal" vote recorders. The majority of this error was a result of vote recorders that used Template 1.

Table 3—Error rates for Test 5, the normal vote recorders (60% Template 1, 40% Template 2)

Machine #	Template #	"0"	"1"	"2"	"3"	% Unsuccessful Votes		Key to Colors
131	1	0	1672	56	72	7.11		
7	1	0	1709	45	46	5.06		errors > 3.00%
26	1	0	1730	21	49	3.89		1.00%< errors < 3.00%
191	1	0	1732	33	35	3.78		errors < 1.00%
17	1	18	1746	13	23	3.00		
66	1	0	1749	17	34	2.83		
155	1	0	1756	11	33	2.44		
83	1	0	1765	15	20	1.94		
167	1	0	1766	16	18	1.89		
42	1	0	1771	13	16	1.61		
36	1	0	1782	7	11	1.00		
62	1	0	1786	9	5	0.78		
97	2	0	1786	12	2	0.78		
95	1	0	1787	6	7	0.72		
75	1	0	1789	6	5	0.61		
107	2	0	1792	4	4	0.44		
194	2	0	1796	1	3	0.22		
116	1	0	1797	2	1	0.17		
61	2	0	1798	1	1	0.11		
27	2	0	1799	0	1	0.06		
46	2	0	1799	1	0	0.06		
19	2	0	1800	0	0	0.00		
28	2	0	1800	0	0	0.00		
108	2	0	1800	0	0	0.00		
132	2	0	1800	0	0	0.00		
Average			**1772**	**12**	**15**	**1.54**		

Table 3 can be compared to Table 4 where only suspect vote recorders were tested in Test 7. Coincidentally, all of the vote recorders used in Test 7, the suspect vote recorders, used Template 1. There were no vote recorders using Template 2 that consistently yielded high error rates. In Test 7, there were only five vote recorders that showed error rates less than

1.0%, and the overall error rate for the Test 7 vote recorders was about 4.0%, which is approximate 2.5% greater than the vote recorders used in Test 5.

Table 4—Error rates for Test 7, the suspect vote recorders (all Template 1)

Machine #	"0"	"1"	"2"	"3"	Total votes	% Unsuccessful Votes		Key to colors
59	7	1618	61	114	1800	10.11		errors > 3.00%
71	1	1632	43	124	1800	9.33		1.00%< errors < 3.00%
15	1	1642	78	79	1800	8.78		errors < 1.00%
13	0	1664	26	110	1800	7.56		
58	0	1678	35	87	1800	6.78		
16	0	1696	29	75	1800	5.78		
9	0	1712	35	53	1800	4.89		
128	0	1726	54	20	1800	4.11		
135	0	1734	36	30	1800	3.67		
29	0	1744	20	36	1800	3.11		
182	0	1752	17	31	1800	2.67		**TEMPLATE #1**
84	0	1768	11	21	1800	1.78		
147	0	1771	9	20	1800	1.61		
130	0	1783	8	9	1800	0.94		
72	0	1786	5	9	1800	0.78		
141	0	1790	4	6	1800	0.56		
63	0	1796	2	2	1800	0.22		
187	0	1799	1	0	1800	0.06		
Average		**1727**	**26**	**46**	**1800**	**4.04**		

Table 5 shows the increase in voter error between Test 5 and 7. This demonstrates that all voters in these tests consistently showed higher error rates when using the suspect vote recorders in Test 7 as compare to the normal vote recorders in Test 5. Since Test 7 included only Template 1, and since the vote recorders used for Test 7 were chosen by a systematic investigation, the conclusion is that there is a high degree of correlation between high error rate and the use of Template 1.

Table 5—Percent increase in voter error from Test 5 to Test 7

Voter	% Increase in error
"03"	3.3
"04"	2.5
"05"	3.1
"06"	2.2
"08"	1.4

The physical defects detailed in the attachment "Vote Recorder Inventory 070201" were also examined to determine if a particular defect correlated with the observation of high error rate. The defects in the vote recorders used in Test 5 and 7 are detailed in the attachments "Vote Recorder Inventory 070201, test 5" and "Vote Recorder Inventory 070201, test 7". These attachments show that 12% of the vote recorders used Test 5 (the "normal" vote recorders)

had some sort of observable mild defect, such as a damaged ballot book or rubber punch base. Coincidently, all of the vote recorders with defects used Template 1. Forty-eight percent of the vote recorders in Test 5 were from the "suspect" set of vote recorders, as denoted by the Chicago Board of Election Commissioners, which appears to be representatively balanced since 46% of the total (219) vote recorders were denoted as "suspect".

Inspection of the vote recorders from Test 7 (the "suspect" vote recorders) showed a higher instance of observable defects, but not so high that they were seen to correlate with the higher error rate. Twenty-two percent of the vote recorders in Test 7 had some physical defect, noting that the physical defects exclude subjective observations such as "more difficult to insert or remove card", which was also noted during the inspection. Of the 22%, only one or two of each of the 18 vote recorders used in the test had the same type of physical defect. In most cases the percentage of defects for each type of defect in Test 7 were only a few percentage points higher than the overall percentages for those of the entire set of 219 vote recorders. Coincidently, all of the vote recorders in Test 7 used Template 1 and 94% of the vote recorders in Test 7 came from the vote recorders originally denoted as "suspect" by the Chicago Board of Election Commissioners. These comparisons show that the physical anomalies of the vote recorders, other than the manufacturing defects in the templates, do appear to correlate with high error rates.

In Test 6, 100 ballots were punched while the depth of the template/ballot tandem into the vote recorder slot was changed. When a card is inserted in to the vote recorder, the bottom of the card reaches the "foot" on the bottom of the template. At this point, the unpunched chads on the ballot should line up with the holes on the template. As more force is applied to the ballot, the ballot card and the template slide in tandem downward against the pressure of the template spring until the holes in the ballot card can be placed over the pins in the vote recorder. When the card is securely resting on the pins, the chads, template holes and ballot book holes should all be aligned so that the stylus can be inserted easily through the assembly. This allows the chad to be removed completely and accurately. In this Test 6, the position of the template/ballot assembly was varied to determine the misalignment that would result in a high probability of error. This type of error could occur if the holes in the template were incorrectly manufactured or if the voter did not line up the ballot correctly on the vote recorder pins.

Figure 2 shows the results of Test 6. The zero value on the horizontal axis represents the point where a ballot and template would typically come to rest if the card were correctly inserted and placed on the alignment pins. The vote recorder was IIT028, which was within manufacturer's tolerance so that the ballot card and the template were aligned. As a result, the only variable that was changing was the translation of the template/ballot pair within the vote recorder. The data in Figure 2 show that at about 0.020" (20 thousandths of an inch), the estimated voting error due to the machine has reached 5%. By 0.040", about ½ the height of a chad, the probably that a correct punch will occur is quickly approaching 0%. On the graph, there are three important notations. First, the ½ chad height, which is the theoretical limit for getting a correct punch, that is, if the template shifts more than ½ a chad, the voter will be punching the ballot card and/or the template frame and not a chad. Secondly, the offset error for Template 1, as measured in Test 8 and 9, which represents the displacement, or misalignment, of the template hole with the chad and the ballot book hole. It can be seen that although this displacement should not produce high errors by itself, it is about half way to

the 5% error point so that any slight inaccuracies with the voter's method for inserting the card and punching the card could result in higher error rates. This probably would be greatly reduced if the templates were within manufacturer's specifications shown in the figure as +/- .002".

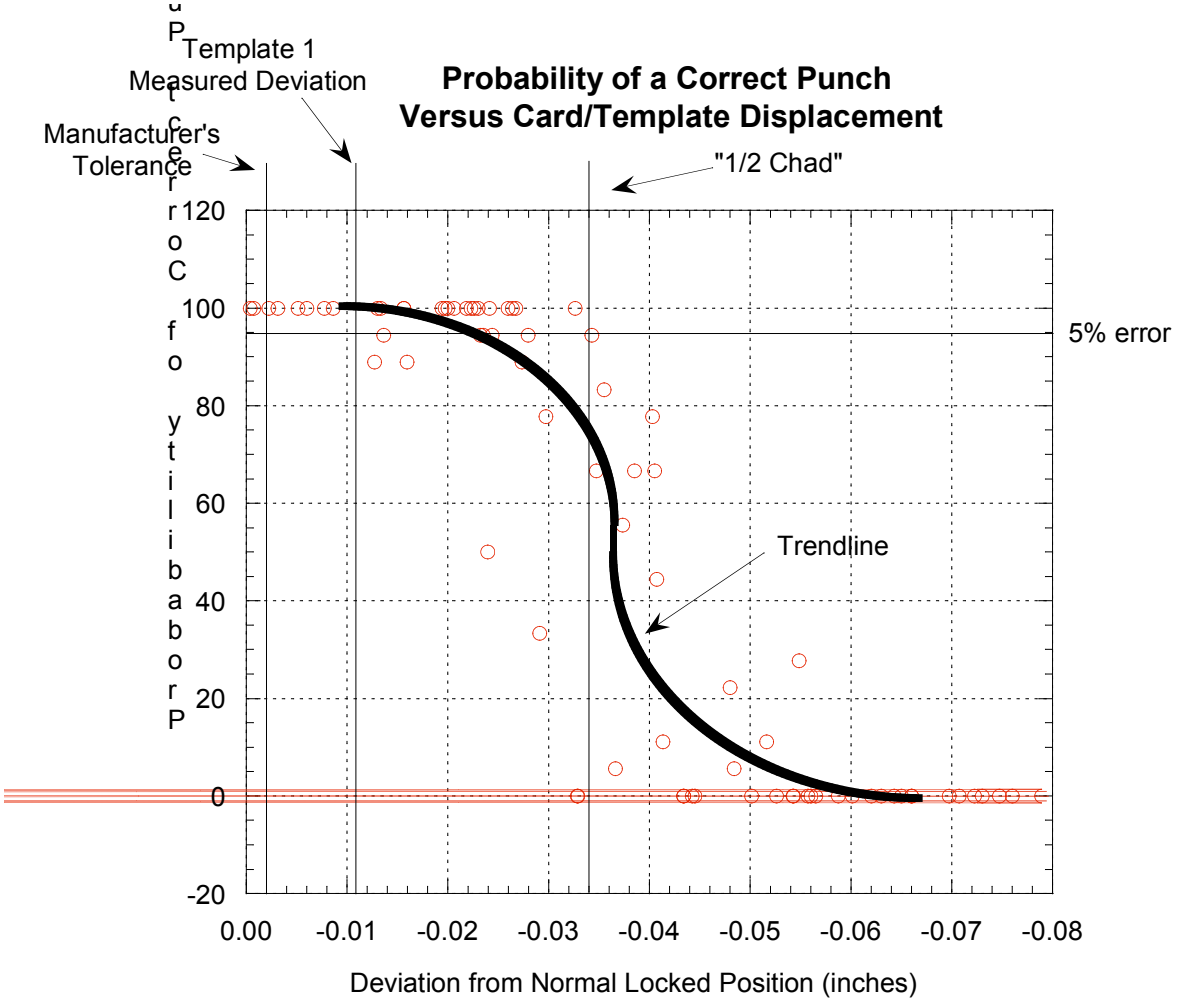

Figure 2 – Probability of a correct punch as a function of template displacement.

Measurements from Tests 8 and 9 characterized the manufacturing accuracy of Template 1 and Template 2. It was found that **both Templates had holes located lower than those specified on the original drawing**, however Template 1 had slightly more than double the error of Template 2. Table 6 shows the average misalignment of the hole as compared to the drawing. Note that the negative sign indicates that the hole was lower than that specified on the drawing. The original tolerances noted in the drawing were 0.002". Figure 3 shows a photograph from the Mitutoyo Quick Scope of a Template 1 hole showing that the template hole is lower than specified on the drawing. The cross hairs in Figure 3 denote the location of the hole in the original drawing.

Table 6 – Vote recorder template hole misalignment.

Template	Average Hole Deviation (Inches)	Standard Deviation (Inches)
1	-0.011	0.002
2	-0.005	0.002

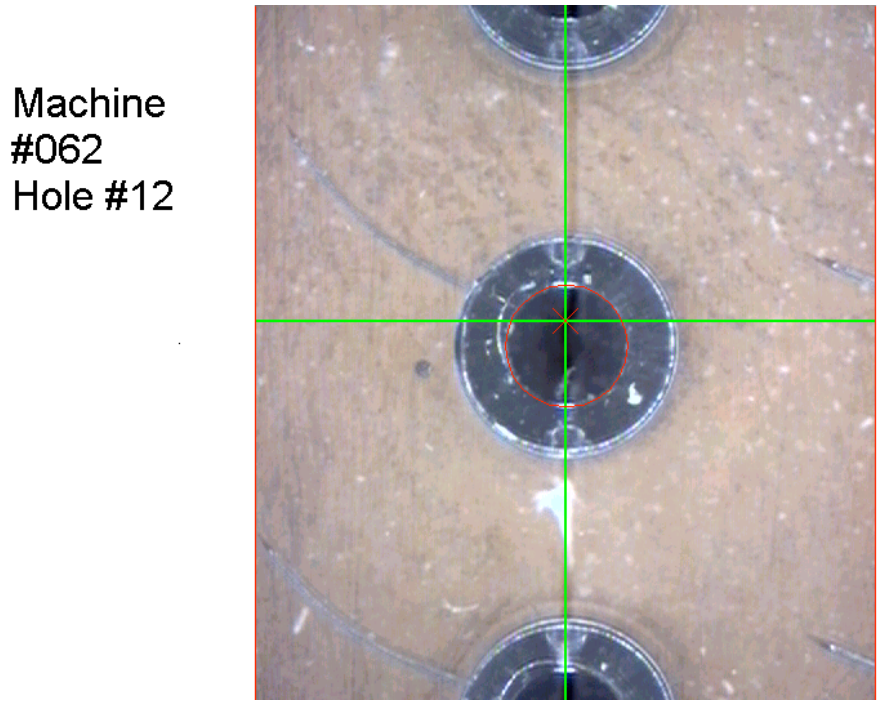

Figure 3 – Photograph of a Template 1 hole showing that the template hole is lower than specified on the drawing.

The voting percentages from all of the tests in this study are summarized in Table 7, separated in to the Template 1 and Template 2 vote recorders. Table 7 shows that the errors in vote recorders using Template 2 are approximately 8 times less than those using Template 1.

Table 7—Summary of all votes for all tests

	All Vote Recorders	Template 1 Vote Recorders	Template 2 Vote Recorders
Total Votes Attempted	224533	152796	71737
Total Successful Votes	220522	148986	71536
% Unsuccessful	1.8%	2.5%	0.3%

The data in Table 7 were statistically analyzed to determine the degree to which the unsuccessful and successful rates were different for Templates 1 and 2. Rearranging the data, Table 7 can be rewritten as Table 8.

Table 8—Statistical analysis of template error

	Unsuccessful	**Successful**	**Total**
Template 1	3810 (2729.51)[a]	148986 (150066.49)[a]	152796
Template 2	201 (1281.49)[a]	71536 (70455.51)[a]	71737
Total	4011	220522	224533

a = Expected value

For ease of interpretation, the odds ratio (θ) statistic and the associated significance test were used in assessing the difference in the two templates (Agresti, 1984). The observed odds ratio can be expressed as:

$$\text{Observed Odds ratio} = \hat{\theta} = \frac{(3810)(71536)}{(201)(148986)} = 9.10.$$

This value of 9.10 means that the odds of getting an unsuccessful vote is 9.1 times higher for Template 1 than it is for Template 2. The expected value of log ($\hat{\theta}$) should be zero under the null hypothesis if the unsuccessful rates were the same for both templates. However, in this example,

log ($\hat{\theta}$) = 2.2084.

Based on this value, and the asymptotic standard error of log ($\hat{\theta}$), we can calculate the 99% confidence interval. The standard error can be shown to be

$$\hat{\sigma}_{(\log\hat{\theta})} = \sqrt{\frac{1}{3810} + \frac{1}{148986} + \frac{1}{201} + \frac{1}{71536}} = .0725.$$

The 99% confidence interval for log ($\hat{\theta}$), based on the current data, is $2.2084 \pm (2.576)(.0725)$ or [2.0216, 2.3952]. If this interval were to include the value zero, then it could be said that the error rates for Template 1 and Template 2 would be similar. However, since this interval does not include the zero point, it indicates that the confidence level is greater than 99% that Template 1 results in significantly higher error rates than Template 2. A chi-square test (Agresti, 1984) of the results in Table 8 also leads to the same conclusion.

The results of these tests indicate that both Template 1 and Template 2 were manufactured outside of the specifications and that vote recorders with Template 1 contributed to a much higher error rate than those with Template 2. The template used in the vote recorder seems to be the single largest contributor to higher error rates. It is recommended that, at a minimum, all of the vote recorders using Template 1 be rebuilt to specifications with new or modified templates. A sample of the vote recorders with the new templates should be tested to confirm that the error rates have decreased to the levels of vote recorders that use Template 2.

Final Comments

The recently released report, "Voting: What is, What Could Be" by Caltech and MIT quantified the residual votes as the total of the uncounted ballots plus the unmarked ballots plus the overvote ballots. For the punch card system, the national average for residual votes was 2.5% for the presidential election. One interpretation of this number is that 2.5% represents the total error in the election, that is, it includes the error due to the vote recorder itself, which was studied here. If this is the case, then the vote recorders with Template 1 from Test 7 that showed an average error of about 4% are clearly above and beyond the national average. Again, the 4% error does not include errors attributable to usability.

The Caltech/MIT report also presented that the Illinois residual vote increased from 2.4% in 1996 to 3.9% in 2000. Finally, the report shows that Cook County reported 6.2% residual vote, which was the second worst, just behind West Palm Beach, Florida at 6.4%. This study indicates that part of this increase in residual vote can be attributed to vote recorder itself.

Reference

Agresti, A. (1984). Analysis of ordinal categorical data. New York: Wiley